

Appendix D. Data Quality

Two major determinants of the quality of income data collected in household surveys are the magnitude of missing responses and the accuracy of the responses that are provided. This appendix has been included to supply information concerning nonresponse rates for selected income questions, the average amounts of income reported in the survey or assigned in the imputation of missing responses, and the extent to which the survey figures underestimate numbers of income recipients and amounts of income received.

Nonresponse in this discussion refers to missing responses to specific questions or "items" on the questionnaire. Noninterviews or complete failure to obtain cooperation from any household member have not been considered in this examination of nonresponse rates. Adjustments to account for noninterviews are made by proportionally increasing the survey weights of interviewed households. Missing responses to specific questions are assigned a value in the imputation phase of the data processing operation.

Nonresponse is a very important factor in assessing the quality of survey data. Nonresponses to income questions cannot be considered random since experience has shown that persons with the highest nonresponse rates have reported characteristics such as education levels and occupations that, in general, differ from population averages. The most frequent causes of nonresponse are the inability of the respondent to answer the question because of either a 1) lack of knowledge or 2) refusal to answer. The first reason is

especially important in situations of proxy response when one household member answers questions for another household member not present at the time of the interview. The practice of accepting proxy interviews from household members deemed "qualified" to answer is a standard procedure in the CPS and most other surveys conducted by the Bureau. During the fourth and fifth interview periods of SIPP, about 37 percent of the interviews were taken from proxy respondents.

The magnitude of nonresponse is generally presented in terms of a nonresponse rate computed by dividing the number of nonresponses by the total number of responses that should have been provided. The first two columns of table D-1 show the number of persons with income and nonresponse rate for a selected group of income amount questions from SIPP for the fourth quarter of 1984. Nonresponse rates for the March 1984 CPS based on annual income amount questions are shown in the third column.

The SIPP nonresponse rates ranged from a low of about 6 percent for Aid to Families with Dependent Children to about 16 percent for self-employment income. These rates were computed by dividing the number of persons with missing responses on the amount received by the total number with either a missing or reported amount for that income type.

The SIPP nonresponse rates for fourth-quarter average monthly amounts contrast sharply with the higher nonresponse rates encountered in the March CPS. The rates for the CPS ranged from a low of 14 percent for food stamp

Table D-1. Persons Nonresponse Rates for SIPP and the March 1985 CPS, and Median Monthly Amounts Reported and Imputed, for Selected Income Types

Income type	SIPP 1984 fourth quarter monthly average		March 1985 CPS nonresponse rate for amounts received	SIPP median monthly average amounts	
	Number with income (thousands)	Nonresponse rate for amounts received		Reported	Imputed
Wage or salary	97,951	7.6	18.9	\$1,151	\$1,273
Self-employment income	8,454	16.1	26.5	884	1,136
Federal Supplemental Security income	3,568	8.4	19.9	212	222
Social Security Income	32,604	12.3	21.9	402	409
Aid to Families With Dependent Children	2,913	5.5	16.0	285	272
Unemployment compensation	2,462	12.7	21.8	366	387
Company or union pensions	8,108	14.7	24.0	236	243
Food stamp allotment	6,523	6.6	13.7	101	90
Veterans' compensation or pensions	3,558	13.5	18.3	128	104

allotments and to 27 percent for self-employment income. The major emphasis given to complete and accurate income information in SIPP and 4-month recall period are two factors that have contributed to the significantly lower nonresponse rates in the SIPP. It should be noted that while the SIPP item nonresponse rates are lower than those in the March CPS, the SIPP household noninterview rate is significantly higher. The March 1984 CPS noninterview rate was about 5 percent compared with the SIPP rate of about 16 percent. The noninterview rate for SIPP was initially about 6 percent but has increased with repeated visits to the sample households.

Nonresponses are assigned values prior to producing estimates from the survey data. The procedure used to assign or impute responses for missing data for SIPP are of a type commonly referred to as a "hot deck" imputation method. This process assigns values reported in the survey by respondents to nonrespondents. The respondent from whom the value is taken is termed the "donor." Values from donors are stored in a matrix defined by demographic and economic data available for both donors and nonrespondents. Each cell of the matrix defines a unique combination of demographic and economic characteristics. For example, the imputation of an amount for monthly wage and salary income is based on eight different variables. These were 1) occupation, 2) sex, 3) age, 4) race, 5) educational attainment, 6) weeks worked, 7) usual hours worked per week, and 8) place of residence.

The last two columns in table D-1 compare median reported and imputed income amounts for SIPP monthly averages, fourth quarter 1984. The differences between reported and imputed median amounts were statistically significant at the 95-percent confidence level for wage and salary income, and self-employment income.

The second important determinant of data quality and probably the one examined most closely by users of the income data collected in household surveys is the accuracy of reported (and imputed) amounts. In general, household surveys have a tendency to underestimate the number of persons receiving income and the average amount received. These problems result for a variety of reasons including ran-

dom response error, misreporting of sources of income, failure to report the receipt of income from a specified source, and failure to report the full amount received. The net effect of these kinds of problems is, for most income types, underestimation or underreporting of income amounts. The extent of underreporting is measured by comparing survey estimates with independently derived estimates, usually based on administrative data that are, generally, more reliable than the estimates derived from the survey. It should be noted that the independent estimates are subject to errors themselves. In addition, independent estimates do not reflect income attributable to the "underground" economy, some of which may be reported in the survey.

Table D-2 contains comparisons of SIPP estimates of the number of persons receiving specific income sources with independent estimates derived from various administrative sources. Table D-3 shows similar comparisons based not on the number of recipients but on the aggregate amount of income received. Data in both of these tables are preliminary and subject to revision.

The comparisons in table D-2 are limited to some of the major transfer programs for which administrative data are available for the October-December 1984 period. Adjustment factors were applied to these administrative figures in order to arrive at the independent estimates for the SIPP noninstitutional population eligible for interview. The adjustment factors used were based on procedures developed by Mathematical Policy Research, Inc., for deriving independent estimates for the 1979 ISDP research panel.

Survey underestimates of income recipients ranged from about 19 percent for Aid to Families with Dependent Children to about 2 percent for Federal Supplemental Security recipients.

The underreporting for Aid to Families with Dependent Children is related to misclassification of this income type as other types of public assistance or welfare. A total of 933,000 persons reported receiving general assistance and 214,000 reported receiving other types of welfare payments for the fourth quarter. A significant number of these cases are actually payments from the Aid to Families with Dependent

Table D-2. Comparison of Estimated Numbers of Income Recipients, for Selected Income Types, Fourth Quarter 1984: SIPP vs. Independently Derived Estimates

(Numbers in thousands)

Income type	Monthly average recipients		SIPP as a percent of independent estimate
	SIPP estimate	Independent estimate	
Federal Supplemental Security Income	3,568	3,637	98.1
Social Security income	32,604	33,438	97.5
Aid to Families With Dependent Children ¹	2,913	3,609	80.7
Unemployment compensation	2,462	2,590	95.1
Food stamp allotment	18,181	19,830	91.7
Veterans' compensation or pensions ¹	3,558	3,815	93.3

¹Excludes dependents covered by payments.

Table D-3. Comparison of Estimated Aggregate Income Amounts Received, for Selected Income Types, Fourth Quarter 1984: SIPP vs. Independently Derived Estimates

(Monthly averages. Figures in millions of dollars)

Income type	SIPP estimate	Independent estimate	SIPP as a percent of independent estimate
Wage or salary	143,199	151,507	94.5
Self-employment income	17,079	(NA)	(X)
Federal Supplemental Security income	785	791	99.2
Social Security income	13,461	13,247	101.6
Aid to Families With Dependent Children	945	1,200	78.8
Unemployment compensation	978	969	100.9
Food stamp allotment	749	896	83.6
Veterans' compensation or pensions	827	1,037	79.7

NA Not available.

X Not applicable.

Children program. This particular problem was also encountered and documented in the developmental ISDP.

Table D-3 provides comparisons of SIPP and independent estimates of the aggregate amount of income received for the total noninstitutional population for the fourth quarter of

1984. Nonseasonally adjusted, monthly independent estimates for wage and salary income is not available. The estimate shown for wage and salary is based on Bureau of Economic Analysis seasonally adjusted, annual rate estimates for the fourth quarter divided by 12. Other independent

Table D-4. Persons 15 Years Old and Over Receiving Income and Aggregate Amount Received, for Selected Income Types: Monthly Average, Fourth Quarter 1984

Income type	Total		Nonfarm	
	Number with income (thousands)	Aggregate amount (millions of dollars)	Number with income (thousands)	Aggregate amount (millions of dollars)
Wage or salary	97,951	143,199	96,296	141,084
Self-employment income	8,454	17,079	7,758	13,963
Social Security income	32,604	13,461	31,666	13,093
Federal Supplemental Security Income	3,568	785	3,527	780
State unemployment compensation	2,462	978	2,422	964
Veterans' compensation or pensions	3,558	827	3,493	815
Workers' compensation	637	493	632	492
Aid to Families with Dependent Children	2,913	945	2,889	941
General assistance	933	179	933	179
Child support	3,252	808	3,238	806
Alimony	432	238	432	238
Company or union pensions	8,108	2,854	7,996	2,811
Federal employee pensions	1,851	1,728	1,822	1,703
U.S. military pensions	1,314	1,335	1,311	1,332
State government employee pensions	2,177	1,145	2,106	1,107
Local government employee pensions	871	468	855	455
Estates and trusts	323	599	309	565
Income from paid-up life insurance or other annuities	695	133	675	129
Money from relatives or friends	1,244	458	1,238	457
Interest income from regular savings accounts, money market deposit accounts, certificates of deposit or other savings certificates, and interest-bearing checking accounts	102,292	6,972	99,548	6,721
Interest income from money market funds, U.S. Government securities, municipal or corporate bonds, and other interest-bearing assets	11,075	1,940	10,831	1,868
Interest on mortgages	3,592	1,372	3,468	1,352
Dividends	20,169	3,704	19,721	3,644
Net rental income	9,935	1,192	9,439	1,120
Income from royalties and other financial investments	2,888	2,102	2,720	2,050

estimates shown in tables D-2 and D-3 are based on various sources including the Social Security Bulletin and unpublished figures from the Department of Health and Human Services, the Department of Agriculture, and the Veterans' Administration.

In most cases the comparisons in table D-3 on aggregate amounts for the fourth quarter parallel the figures in table D-2 for estimated number of recipients. The comparison for wage and salary income is difficult to interpret because the independent estimate is seasonally adjusted. A monthly independent estimate for self-employment income is not

available because the self-employment income estimates are based on different concepts. The SIPP figure is based on the "salary" and other income received from the business by the owners. More refined comparisons between SIPP estimates and estimates derived from independent sources will be made in future reports.

Table D-4 shows the monthly averages for the number of income recipients and aggregate amounts of income received for the fourth quarter 1984 for the total population and the nonfarm population. Most of the largest sources of income have been included in this table.